Environmental Modelling & Software xxx (2014) 1–10

Contents lists available at ScienceDirect

Environmental Modelling & Software

journal homepage: www.elsevier.com/locate/envsoft

Methods for uncertainty propagation in life cycle assessment

E.A. Groen [a,*], R. Heijungs [b,c], E.A.M. Bokkers [a], I.J.M. de Boer [a]

[a] Animal Production Systems Group, Wageningen University, PO Box 338, 6700 AH, Wageningen, the Netherlands
[b] Institute of Environmental Sciences, Leiden University, PO Box 9518, 2300 RA, Leiden, the Netherlands
[c] Department of Econometrics and Operations Research, VU University Amsterdam, 1081 HV, Amsterdam, the Netherlands

ARTICLE INFO

Article history:
Received 22 October 2013
Received in revised form
15 September 2014
Accepted 9 October 2014
Available online xxx

Keywords:
Latin hypercube sampling
Quasi Monte Carlo sampling
Fuzzy interval arithmetic
Analytical uncertainty propagation
Fisheries

ABSTRACT

Life cycle assessment (LCA) calculates the environmental impact of a product over its entire life cycle. Uncertainty analysis is an important aspect in LCA, and is usually performed using Monte Carlo sampling. In this study, Monte Carlo sampling, Latin hypercube sampling, quasi Monte Carlo sampling, analytical uncertainty propagation and fuzzy interval arithmetic were compared based on e.g. convergence rate and output statistics. Each method was tested on three LCA case studies, which differed in size and behaviour. Uncertainty propagation in LCA using a sampling method leads to more (directly) usable information compared to fuzzy interval arithmetic or analytical uncertainty propagation. Latin hypercube and quasi Monte Carlo sampling provide more accuracy in determining the sample mean than Monte Carlo sampling and can even converge faster than Monte Carlo sampling for some of the case studies discussed in this paper.

Software and data availability

All modelling done in this paper is done in MATLAB®; the code and data can be forwarded by the first author upon request. In addition, software for life cycle assessment calculations based on matrix representation is available at CML, and was developed by Reinout Heijungs, Leiden University. The software CMLCA can be downloaded free of charge from: www.cmlca.eu, programming language Delphi, system requirement Windows XP or higher (32 bits), no special hardware requirements.

1. Introduction

Life cycle assessment (LCA) is an established method to calculate the environmental impact of a product over its entire life cycle (Curran, 2012). It has also been applied in areas such as business strategy, product innovation, policy development and eco-labelling (Cooper and Fava, 2006). It has been used to quantify environmental emissions of production systems (e.g. Whittaker et al., 2013), and to make decisions about potential options to reduce the environmental impact of products (Carvalho et al., 2012; Levis et al., 2013; Tillman, 2000). According to the ISO 14040 standardized framework for (environmental) life cycle assessment, an LCA consists of four phases: goal and scope definition, inventory analysis, impact assessment and interpretation. The goal and scope phase includes definition of the boundary of the product life cycle, the functional unit (main output of the system) and allocation method. In the inventory analysis, data about resource use and emissions related to each production process are collected and are translated to environmental impacts. During the final phase the results are presented and interpreted.

During the inventory analysis, data are collected from different systems along the chain. These data can be highly variable, especially if they originate from systems that depend on weather conditions, like agriculture (Brentrup et al., 2000). Results of an LCA, therefore, are uncertain due to lack of knowledge about the true value of its model parameters (Björklund, 2002; Heijungs and Huijbregts, 2004). LCAs that demonstrate a single point value as their result, overlook the range of possible realizations of output data, and could therefore be misleading (Björklund, 2002), or might provide a false sense of accuracy (De Koning et al., 2010). At present, a standardized definition for uncertainty in LCA (Björklund, 2002; Heijungs and Huijbregts, 2004; Finnveden et al., 2009) and a standardized methodology how to propagate uncertainties are missing. Application of uncertainty propagation in LCA might be hampered due to long calculation time (Ciroth et al., 2004), caused by the large amount of data elements (i.e. input parameters); lack of consensus about relevant applicable methods or missing information on the required descriptive statistics of input parameters

* Corresponding author. Tel.: +31 (0)317 484626; fax: +31 (0)317 485550.
E-mail addresses: evelyne.groen@wur.nl (E.A. Groen), r.heijungs@vu.nl (R. Heijungs), eddie.bokkers@wur.nl (E.A.M. Bokkers), imke.deboer@wur.nl (I.J.M. de Boer).

http://dx.doi.org/10.1016/j.envsoft.2014.10.006
1364-8152/

(Björklund, 2002). Incorporating uncertainty, however, increases reliability of results and improves decision making.

Various types and sources of uncertainty can be distinguished (Björklund, 2002; Heijungs and Huijbregts, 2004; Finnveden et al., 2009). In this paper, we focus on parameter uncertainty, following the definition of Björklund (2002) and Heijungs and Huijbregts (2004). Parameter uncertainty includes e.g. inaccuracy of measurements, erroneous data, incomplete data, round-off errors and (natural) variability. A comprehensive paper of Lloyd and Ries (2007) showed the broad scope of techniques and methods available to researchers in LCA towards propagation of (parameter) uncertainty. Uncertainty propagation is currently dominated in LCA literature by one type of method (i.e. Monte Carlo sampling), but there are several other methods on propagating uncertainties (Lloyd and Ries, 2007). These methods would be better at coping with the large size of LCA inventories and considerably reduce computational effort (i.e. calculation time and memory usage) (Heijungs, 2010), or can be applied when limited knowledge about the input uncertainties is present (Heijungs et al., 2005).

So far, these methods of uncertainty propagation have not been compared in a consistent manner. Methods examined in this paper were selected from Lloyd and Ries (2007), because they are most commonly applied. The following methods were selected: Monte Carlo sampling (MCS) and (standard) Latin hypercube sampling (LHS), analytical uncertainty propagation (AUP) and fuzzy interval arithmetic (FIA) (Lloyd and Ries, 2007). In addition, we also selected another type of sampling method: (randomized) quasi Monte Carlo sampling (QMCS) that showed promising results (e.g. Subramanyan et al. (2008) in LCA and Tarantola et al. (2012) more in general).

The performance of these uncertainty propagation methods may depend on size, behaviour of the case study (linear or non-linear) and properties of the input parameters (Saltelli et al., 2010; Tarantola et al., 2012). To illustrate the performance of the methods, we therefore examined three case studies that differed in size and behaviour, and type of distribution function and dispersion of the input parameters. Two small artificial case studies were selected that differed with respect to their behaviour and a bigger linear case study representing an existing production system of a north-east Atlantic fishery.

This study aims to (1) give more insight into Monte Carlo sampling, Latin hypercube sampling, quasi Monte Carlo sampling, fuzzy interval arithmetic and analytical uncertainty propagation, and (2) compare the uncertainty propagation methods to show the performance, advantages and disadvantages of each method used when applied in LCA. We start with a recap of matrix representation in LCA, followed by an introduction to each of the five uncertainty propagation methods. Subsequently, all five methodologies are applied to the case studies and their results are compared on different criteria.

2. Materials and methods

2.1. Preliminary on notation and terminology

In order to calculate the environmental impact (**g**) corresponding to a functional unit (**f**), a model of the following form is constructed (Heijungs and Suh, 2002):

$$\mathbf{g} = \mathbf{B}\mathbf{A}^{-1}\mathbf{f} \tag{1}$$

The (square) technology matrix **A** represents production processes which are given in each column, the rows represent product flows. The matrix is constructed in such a way that e.g. in the first column electricity is produced that is subsequently used for fuel production in the second column. This results in a large matrix with interlinked production processes, which is scaled to produce the amount given by the functional unit vector **f** (the scaling vector **s** equals: $\mathbf{s} = \mathbf{A}^{-1}\mathbf{f}$). The intervention matrix **B** represents the input of raw materials and output of emissions corresponding to each production process of the technology matrix **A**. In most LCAs a readily available database is used for **A**, like the ecoinvent v2.2 database (ecoinvent, 2010), which is currently made up out of 4087 processes. The size of this square **A**-matrix, therefore, is large, although many elements are zero. When using readily available databases like the ecoinvent database, as is quite common in LCA, the size of **A** causes calculation of the environmental impact to be of high computational effort. This is due to the calculation of the inverse of **A**, which slows down the calculation. Especially when performing an uncertainty analysis using a stochastic approach and large amounts of simulations are required. This illustrates the exploration of low demanding sampling methods to decrease calculation time and memory usage.

2.2. Uncertainty propagation methods

In the next section five uncertainty propagation methods will be discussed: MCS, LHS, QMCS, AUP and FIA. Subsequently, a description of the case studies and their model parameters are given, followed by a description of how the methods are compared.

2.3. Monte Carlo sampling

Uncertainty propagation using a sampling approach started with the development of MCS in 1949 (Metropolis and Ulam, 1949), quickly becoming widespread (Burmaster and Anderson, 1994; Helton et al., 2006). MCS consists of drawing (pseudo-) random numbers from a set of input parameters (k) with known distribution functions to obtain the sampled distribution of the output parameter (Helton et al., 2006). While MCS might be computationally demanding, it has the advantage that it can be used to compare output statistics like the mean or parameter of dispersion between two studies. A disadvantage is the long calculation time due to a large number of simulations that have to be performed before the output uncertainty can be determined.

The standard error of the mean (SEM) can be used as a measure of the convergence rate of the sampling method. The SEM describes how much variation is expected around the sample mean for a specific sample size. In theory, the standard error of the mean of MCS convergences as $\mathcal{O}(s/\sqrt{N})$, where s is the standard deviation of the sample and N is the sample size. This means that the convergence rate does not depend on the amount of input parameters.

In order to apply MCS, the following input is required: the type of distribution function (e.g. log-normal, triangular) of each input parameter, a central value (such as the mean (μ)) and a parameter of dispersion (π), such as the standard deviation (σ) of the probability density function in case of normal distributions. The result of applying this method is a sampled distribution of **g** (Table 1).

2.4. Latin hypercube sampling

LHS is a variant of Monte Carlo sampling employing a stratified sampling approach. The distribution functions of the input variables are divided in equally probable (i.e. stratified) subgroups from which random numbers are drawn, which are subsequently combined at random (Helton et al., 2006). The idea behind LHS is that input parameters are sampled more uniformly than (psuedo-) random sampling, resulting in a faster convergence rate. LHS has been suggested, therefore, as a promising alternative to MCS for uncertainty propagation in LCA by Heijungs (2010) and Peters (2007), and it is used in LCA amongst others by Basson and Petrie (2007); De Koning et al. (2010); Geisler et al. (2005) and Huijbregts et al. (2000). Many improvements have been made to the original LHS design, such as maximin LHS (Morris and Mitchell, 1995) and Latin supercube sampling (Tarantola et al., 2012), but we will only focus on the original design, referred to in this paper as standard LHS. In order to apply this method, similar type of information as for MCS is required (Table 1).

2.5. Quasi Monte Carlo sampling

The QMCS procedure is similar to MCS, but uses quasi-random numbers to sample from the distribution functions (Sobol', 1967; Saltelli et al., 2010). Quasi-random numbers are deterministic numbers that are equally distributed for a

Table 1
Descriptive statistics and additional requirements for implementation of Monte Carlo sampling (MCS), Latin hypercube sampling (LHS), quasi Monte Carlo sampling (QMCS), Fuzzy interval arithmetic (FIA) and analytical uncertainty propagation (AUP) in life cycle assessment.

Method	Inputs	Outputs	Additional requirements
MCS	pdf[a], μ, π	e.g. $\bar{x}$, s	(pseudo-) Random sample generator
LHS	pdf, μ, π	e.g. $\bar{x}$, s	Stratified sampling approach, (pseudo-) random sample generator
QMCS	pdf, μ, π	e.g. $\bar{x}$, s	Quasi-random numbers, randomization method
FIA	S, v_c, $\delta^{\pm}$	$\bar{v}_c$, $\bar{\delta}^{\pm}$	Construction of a possibility function
AUP	σ	σ	First order Taylor approximation

[a] pdf: probability density function (distribution function) μ: mean;π: parameter of dispersion; $\bar{x}$: mean of the sample; s standard deviation of the sample; S: shape; v_c: core value; $\delta^{\pm}$: upper and lower bound; σ: standard deviation.